Journal of Traumatic Stress, Vol. 16, No. 1, February 2003, pp. 59–66 (© 2003)

Is the Emotional Stroop Paradigm Sensitive to Malingering? A Between-Groups Study With Professional Actors and Actual Trauma Survivors

Todd C. Buckley,[1,2,4] Tara Galovski,[3] Edward B. Blanchard,[3] and Edward J. Hickling[3]

Six professional actors, trained by psychologists and acting coaches to feign PTSD, were covertly enrolled into a treatment outcome study for PTSD with the aim of investigating malingering. During pretreatment assessment, individuals completed an emotional Stroop task. Vocal response latencies to different classes of stimuli were examined for sensitivity to malingering. Actor response latencies were compared to those of 6 nonlitigant PTSD patients and 6 nonanxiety controls. The actor/dissimulation group was able to feign an overall slowing of response latency across stimulus types, similar to the PTSD group. However, they were unable to modulate response latency as a function of stimulus content, a pattern that characterized the PTSD group. The use of information-processing paradigms to detect dissimulation is discussed.

KEY WORDS: malingering; Stroop; PTSD.

Introduction

Since the introduction of PTSD into the formal diagnostic system in 1980 (American Psychiatric Association, 1980), the diagnosis and treatment of the disorder have often been the subject of controversy. This is due in part to the fact that PTSD, by definition, can only result from clearly identified stressors such as natural disasters (e.g. earthquakes) or events that often result from negligence or deliberate acts of a second party (e.g., motor vehicle accidents [MVA], sexual assault; Blanchard & Hickling, 1997; Rothbaum, Foa, Riggs, Murdock, & Walsh, 1992). Therefore, PTSD holds a relatively unique role in the psychiatric nomenclature in that the event associated with onset of disorder can often be directly attributable to the fault of a party other than the individual receiving the diagnosis. For this reason, the PTSD diagnosis is one that is often litigated and used by prosecuting parties to establish the basis of pain and suffering settlements (Melton, Petrila, Poythress, & Slobogin, 1997).

Given the secondary gain contingencies that operate in court cases where one is the victim of a second party's negligence or deliberate action, much concern has been raised about the possibility of plaintiffs exaggerating PTSD symptoms or outright faking them (Fairbank, McCaffrey, & Keane, 1985). The Department of Veteran's Affairs has been faced with similar issues when trying to determine the legitimacy of PTSD claims during compensation and pension evaluations for veterans exposed to wartime atrocities (Freuh, Cahill, Hamlin, Gold, & Hamner, 2000). It is not uncommon for reports to appear in the literature, which suggest that individuals who are claiming a PTSD diagnosis in the presence of secondary gain contingencies overreport symptoms (for a review, see Freuh et al., 2000).

[1]National Center for PTSD, VA Boston Healthcare System, Boston, Massachusetts.
[2]Boston University School of Medicine, Boston, Massachusetts.
[3]Center for Stress and Anxiety Disorders, University at Albany – State University of New York, Albany, New York.
[4]To whom correspondence should be addressed at National Center for PTSD (116B-2), VA Boston Healthcare System, 150 South Huntington Avenue, Boston, Massachusetts 02131-4817; e-mail: todd.buckley@med.va.gov.

0894-9867/03/0200-0059/1

In an attempt to determine the validity of patient claims of disorder, clinicians have only a limited number of diagnostic methods to which they can turn. The most common method of detecting deviant response sets is through the use of empirically derived validity/clinical indices from self-report instruments such as the MMPI/MMPI-2 (Minnesota Multiphasic Personality Inventory, Berry, Baer, & Harris, 1991; Lyons, Caddell, Pitman, Rawls, & Perrin, 1994). It has been shown, however, that such tests are susceptible to dissimulation because many of the critical items are high in face validity and the nature of deviant responses is easily discernible to the test taker (Frueh & Kinder, 1994; Lyons et al., 1994). It is also the case that the validity indices of such measures often highly correlate with genuine affective distress, thereby complicating the decision point between true pathology and malingering (Frueh et al., 2000).

Others have suggested that more "objective" measures such as physiological responses to trauma-related cues might be an effective adjunctive assessment tool in determining PTSD diagnostic status (Orr & Pitman, 1993; Pitman & Orr, 1993). In a similar vein, it has been proposed that measuring startle reflexes to neutral stimuli such as acoustic probes might also serve as a useful adjunct to assessment (Morgan, 1997). Both of these psychophysiological techniques have demonstrated adequate positive and negative predictive power of diagnostic status in research efforts with traumatized populations (see Blanchard & Buckley, 1999). However, these methods have not been utilized often in clinical settings because of the equipment costs, expertise, and time involved in administering the physiological assessments. Furthermore, research shows that physiological markers are not completely free from dissimulation as non-PTSD veterans instructed to "fake bad" have demonstrated an ability to evince autonomic reactivity similar to that seen in actual PTSD veterans (Gerardi, Blanchard, & Kolb, 1989).

Another set of potentially powerful assessment techniques, which do not rely on the self-report of the patient, are reaction time measures from information-processing tasks (Mogg, Kentish, & Bradley, 1993). Over the past two decades clinical researchers adopted many paradigms that have their origin in experimental–cognitive psychology in order to elucidate the manner in which diagnostic subtypes are characterized by variation in attentional biases according to stimulus class (Mathews & MacLeod, 1994). These methods have provided quantifiable and reliable means of measuring the way in which clinical groups process various classes of stimuli differently than control groups free from psychopathology (Williams, Mathews, & MacLeod, 1996).

Perhaps the most widely used information-processing reaction time task with clinical populations is the emotional analogue of the Stroop paradigm (Williams et al., 1996). During this task, patients are presented with stimuli of varying hedonic valence. The stimuli also vary in semantic relation to the clinical concerns of the patient. It is the task of the participant to name the color of the stimulus as quickly as possible while ignoring the semantic content of the word itself. A very reliable phenomenon across numerous studies with traumatized individuals is that patients with PTSD show delayed vocal response latencies for trauma-relevant stimuli relative to neutral classes of stimuli or threat stimuli that are not trauma related (Buckley, Blanchand, & Neill, 2000; Foa, Feske, Murdock, Kozak, & McCarthy, 1991). Groups free of PTSD do not show this pattern of stimulus-class-specific responding. Such differences in responding across diagnostic groups and stimulus classes have been attributed to differences in attention allocation (see McNally, 1995, for a review of process mechanisms that may account for this effect).

Reaction time tasks such as the Stroop may be sensitive to malingering because the direction and magnitude of a "disordered" response is less readily apparent to the patient/participant relative to self-report or structured interview questions, the "correct" answers of which may be discernable to potential malingerers (Frueh et al., 2000). Despite the vast literature on information-processing tasks and PTSD, we are unaware of any studies that have directly examined their sensitivity to malingering. This paper discribes a study in which six survivors of severe MVAs, who were not involved in litigation and had received a diagnosis of PTSD, participated in a multimodal assessment procedure that included an information-processing reaction time task. Their reaction time measures were compared to the responses of six professional actors who were trained by doctoral-level psychologists and acting coaches to feign a PTSD reaction to a stressor that never occurred. The actor grup went through the same multimodal assessment as the clinical participants. A nonanxiety control group *without* instruction to feign symptoms was also included in the study. All doctoral-level and doctoral student staff involved in the administration of the assessment protocol were unaware that actors had been enrolled in the study as "fake" accident victims.

Method

Parent Study

The six professional actors were covertly enrolled in the pretreatment assessment portion of a large

randomized-controlled-outcome study for cognitive–behavioral treatment of PTSD. Their participation included a comprehensive, multimodal assessment battery consisting of self-report psychometric instruments, structured clinical interviews, open clinical interviews, and a standardized psychophysiological laboratory challenge procedure. The overall purpose of the "actor" portion of the study was to identify (statistically) those variables that are most sensitive to malingered response sets and to examine the ability of clinicians to detect malingering. Thus, the actors were in one capacity confederates in the study and in another also provided research data, whereas the two other groups (MVA survivors and controls) were considered research subjects solely. The clinicians who conducted the interviews on the actors (whom they believed to be accieent victims) also were considered study participants. The IRB of the University at Albany-SUNY approved this deception for the purposes of the study, with the proviso that all clinicians and actors be fully debriefed following the study. The two other groups (MVA survivors and controls) were not deceived and underwent standard informed consent procedures as approved by the local IRB.

The comprehensive methodology and primary results for this study have been published elsewhere (Hickling, Blanchard, Mundy, & Galovski, 2002). In short, the Hickling et al. paper reports on findings from clinical interview data, psychometric instruments, and psychophysiological laboratory assessment for the purposes of detecting malingering. It was found that none of the dissimulators were initialy detected on the basis of spontaneous reporting of the interviewing clinicians, nor was there any indication in the lengthy reports prepared from the initial assessment that the assessors were suspicious of malingering. However, once informed they had been deceived by one of their previous assessment cases, clinicians were able to retroactively identify 50% of the cases with accuracy. Psychometric instruments (e.g., Beck Depression Inventory, BDI) showed no differences that differentiated malingerers from true cases. The psychophysiological reactivity data showed some sensitivity and potential for distinguishing malingerers simulating PTSD symptoms (Hickling et al., 2002). Readers interested in more detailed information about these analyses are referred to that paper specifically.

The current paper focuses on an unpublished data set from a separate laboratory procedure that was not part of the Hickling et al. examination. Therefore, despite the relationship of the two papers to a common parent study, the findings reported herein represent unique data on an information-processing task that are not presented elsewhere. The data were derived from a concurrent portion of the pretreatment assessment, namely, the emotional Stroop paradigm. The rest of the paper is dedicated to that portion of the study only.

Design

The design of this study was a 3 (Group: PTSD, actor, nonanxiety control) × 2 (Word-type: PTSD-related, categorized neutral) factorial with repeated measures on the second factor. The primary dependent variable was vocal response time, measured in milliseconds.

Participants

PTSD Group

The PTSD group consisted of six survivors of severe MVAs who were 6–24 months post-MVA, met criteria for PTSD, and were not involved in litigation as a result of their MVA. This group was recruited through self-referral, response to local media advertising, and referral from the medical community. MVA victims who suffered a closed head injury (evidenced by loss of consciousness) as a result of their MVA were excluded from the study. In addition, MVA survivors meeting criteria for current substance dependence and/or psychotic disorders were also excluded.

Actor Dissimulation Group

A group of six professional actors were taught the diagnostic criteria for *DSM-IV* diagnoses of PTSD and major depression. A licensed doctoral-level psychologist provided this group with basic information about the nature of psychological disorders that frequently follow MVAs, and provided rationale for the need of this type of research. A method acting coach was in attendance during the dissemination of this information and assisted the actors in the portrayal of being an MVA survivor. Accident scenarios were constructed for each actor to utilize in an effort to effectively fake being a psychologically affected survivor of a severe MVA. The actors were also instructed to draw upon their own experiences and improvise as necessary to convince clinic staff that they were eligible for the study. The actors were not provided with information about how individuals with PTSD would respond on a modified Stroop task, but rather, were informed about specific symptom criteria (*DSM-IV*; American Psychiatric Association, 1994) and instructed to stay in role and respond as they thought necessary to present as if they were

suffering from PTSD (based on their knowledge of symptom criteria). The actors were then given the phone number of the clinic and called in "in response to the newspaper advertisement" that was running concurrently with their involvement in the project. A doctoral student (unaware of the simulation nature of the call) conducted a phone screen with the participants and enrolled them in the study protocol. All clinic staff who conducted assessments with the actors were very experienced in assessing MVA survivors. The assessing clinicians revealed no awareness that these six individuals were not actual MVA victims. The PI of the project informed the assessors of the dissimulation nature of these subjects only after the completion of all six assessment protocols.

Nonanxiety Control Group

The nonanxiety control group consisted of six individuals recruited through local media advertising. Subjects with any current anxiety disorder or a history of MVA-related PTSD were excluded from this group. Subjects with a history of serious head injury, current substance dependence, and/or a current psychotic disorder diagnosis were also excluded.

Measures

The Structured Clinical Interview for *DSM-IV* (SCID-I; First, Spitzer, Gibbon, & Williams, 1996) was used to determine the presence/absence of all lifetime and current Axis-I diagnoses for all three groups. The Clinician-Administered PTSD Scale (CAPS; Blake et al., 1997) was administered to both the MVA and actor groups to determine whether subjects met *DSM-IV* diagnostic criteria for PTSD.

To assess the level of PTSD symptoms, the MVA and actor groups were administered the PTSD Checklist (PCL; Weathers, Litz, Herman, Huska, & Keane, 1993) and the Impact of Event Scale (IES; Horowitz, Wilmer, & Alvarez, 1979). All three groups completed the BDI (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) and the State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, & Lushene, 1970) to quantify depressive and anxiety symptoms respectively. The National Adult Reading Test—Revised (NART-R; Blair & Spreen, 1989) was administered to all participants to determine level of verbal ability. We also gathered demographic information including age, gender, and level of educational achievement. Demographic information on the diagnostic subgroups can be found in Table 1.

Table 1. Demographic Characteristics of PTSD, Actor, and Control Groups

	PTSD	Actor	Control
Age	34.7 (7.0)	26.2 (18.5)	28.7 (12.2)
Gender (M/F)	0/6	2/4	1/5
Years of education	13.8 (2.6)	12.8 (0.8)	15.2 (2.3)
BDI	27.5 (11.8)	32.0 (16.5)[a]	2.7 (2.7)
State anxiety	53.5 (11.4)	55.3 (16.5)[a]	28.3 (6.9)
Trait anxiety	58.5 (9.6)	64.2 (11.5)[a]	31 (6.5)
PCL	53.7 (9.5)	56.8 (13.1)[a]	N/A
CAPS score	69.5 (15.9)	63.3 (14.3)[a]	N/A
NART-R	36.5 (15.0)	41.6 (10.0)	35.5 (11.4)

Note. BDI = Beck Depression Inventory; CAPS = Clinician-Administered PTSD Scale; PCL = Posttraumatic Checklist; NART-R = National Adult Reading Test—Revised.
[a]Feigned responses.

Stimulus Materials

In total there were two categories of words, neutral words and PTSD-related threat words. The neutral word stimuli consisted of 16 words (two lists of eight semantically related words; tools and musical instruments). Using semantically related words for the neutral word categories controlled for semantic priming effects. The PTSD-specific threat stimuli included words that have produced delayed vocal response times (relative to neutral words) on modified Stroop tasks with MVA-related PTSD populations (i.e., *highway*, *crash*, etc.; Bryant & Harvey, 1995; Harvey, Bryant, & Rapee, 1996). A one-way analysis of variance (ANOVA) on a frequency of occurrence measure indicated that these word groups did not differ in their frequency of usage in the English language, $F(2, 21) < 1$ (Carroll, Davies, & Richman, 1971). In addition, a one-way ANOVA revealed that the three word groups did not differ in the number of characters per word, $F(2, 21) < 1$. The word stimuli used in this study are available from the first author upon request.

Each word appeared four times, once in each of four colors (green, blue, white, and red), yielding 96 total trials. The words were presented in a fixed-randomized format with the following constraints. No two words from the same category appeared consecutively, nor did a color appear twice on consecutive trials. The order of presentation was determined through use of a random number table.

Apparatus

An IBM computer with a 100 MHz Pentium processor presented the stimuli individually to the subjects and recorded their vocal response latency to name the color of stimuli on each trial (in milliseconds). MEL Professional software (version 2.0d) was used to program

the experiment. The stimuli appeared in 1-in. capitalized block letters in the center of the screen, on a 14-in. Magnavox color monitor. A voice-activated relay connected to a Software Tools 200A-Psychology serial response box was used to detect the onset of vocal responses.

Procedure

When prospective participants contacted the clinic for participation in the study, a brief phone interview was conducted. After a phone screen contact indicated that subjects might be eligible for the study, they were mailed the self-report questionnaires to complete prior to their first appointment for the structured clinical interviews. Prior to beginning the assessment, all participants gave written informed consent. Following the structured interview session, the patients were asked to report back to the clinic approximately 1 week later for Stroop testing. In this way, semantic and emotional priming effects for disorder-specific threat words were minimized (as opposed to running the Stroop task immediately after the clinical interview).

The Stroop task was described as one which tests concentration. Participants were seated in a sound-attenuated room approximately 18 in. from the computer monitor. All participants wore a headset microphone to activate the vocal response relay.

Stimulus words were preceded by a 1-in. white fixation cross, which appeared in the middle of the computer screen at the same spot where the word stimuli appeared. The interstimulus interval was 2 s. Prior to the experimental trials, patients responded to 10 practice trials with words unrelated to the experimental stimuli (e.g., ONE, TWO, THREE). Provided the participants understood the task, the 96 experimental trials were conducted.

Results

Given violations of sphericity on the repeated measure created by small sample size, a split-plot ANOVA analysis was not conducted. Rather, a series of Kruskal–Wallis one-way ANOVAs, which do not require assumptions of normality and are appropriate for small sample data such as this (Heiman, 1996), were utilized to examine the vocal response latencies. Effect size analyses (Cohen's *d*; Cohen, 1988) were also computed for all comparisons so as not to rely strictly on *p* values (Schmidt, 1996). The mean vocal response latencies and standard errors for both trauma-related words and neutral words are presented in Fig. 1 as a function of diagnostic group status.

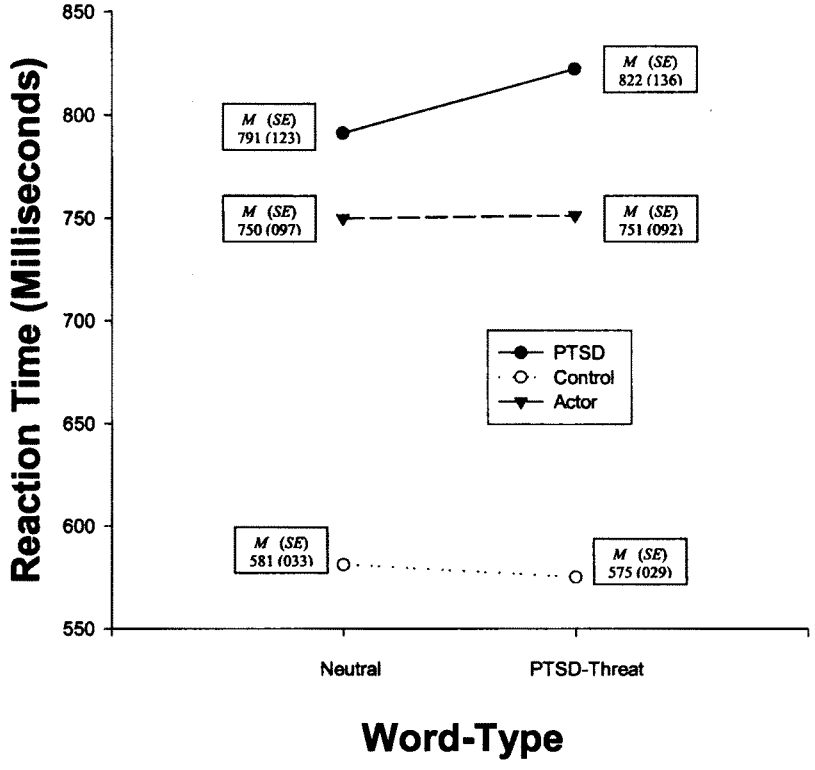

Fig. 1. Mean vocal response latencies across word-type as a function of diagnostic group.

To examine differences in responding as a function of stimulus content, interference scores were computed for each group by subtracting the vocal response latency for neutral words from that of the threat words. The resultant score reflects the magnitude of differences in response latencies across the two stimulus classes (in milliseconds). This method of examining interference scores across stimulus types is very common for both experimental and clinical Stroop methodology (MacLeod, 1991; Williams et al., 1996). For both the nonanxiety and actor groups, there was no increase in vocal response delay in the threat word condition relative to the neutral words (see Fig. 1). The PTSD group, however, showed on average, a 31-ms delay in vocal response time for threat words relative to neutral words. A Kruskal–Wallis one-way ANOVA on these interference scores revealed that the difference between groups was statistically significant, $\chi^2(2, N = 18) = 6.50$, $p < .05$. Follow-up Kruskal–Wallis ANOVAs revealed that the PTSD group showed greater interference than did the control group, $\chi^2(1, N = 12) = 6.56$, $p < .05$, and differences that approached significance ($p = .14$) relative to the actor group, $\chi^2(1, N = 12) = 2.13$. The differences between the actor and control groups were not significant, $\chi^2(1, N = 12) < 1$.

To determine the magnitude of effect between groups, we computed Cohen's d as a measure of effect size (Cohen, 1988). The effect size for the comparison between the PTSD group and the actor dissimulation group was large ($d = 0.92$). The comparison between the PTSD group and the control group was also large ($d = 1.15$). However, the difference between the control group and the actor dissimulation group was considerably smaller ($d = 0.23$). These effect sizes suggest that the actor and control groups were very similar in terms of a lack of differences in response latencies across stimulus class, whereas the PTSD group was markedly different from both the actor and control groups in that respect (they showed slower responses to the PTSD-related stimuli relative to neutral stimuli).

A group main effect for overall reaction time was analyzed by computing a reaction time index for each group collapsed across word-types. A Kruskal–Wallis one-way-ANOVA on this measure indicated that the group main effect just misses conventional levels of statistical significance, $\chi^2(2, N = 18) = 5.37$, $p = .06$. Follow-up ANOVAs reveal that the PTSD and actor groups were comparable in overall response latency, $\chi^2(1, N = 12) < 1$. However, the PTSD group was slower than the control group, $\chi^2(1, N = 12) = 4.33$, $p < .05$, as was the actor group, $\chi^2(1, N = 12) = 3.33$, $p = .06$. Effect size analyses revealed that the comparison between the PTSD and actor groups was modest for overall reaction time ($d = 0.22$), whereas the comparisons between PTSD versus Control ($d = 0.96$) and Actor versus Control ($d = 0.74$) were considerably larger. These effect sizes suggest that, collapsed across stimulus types, the PTSD and actor groups were similar, with both being different from the control group (that is to say, the PTSD and actor groups show an overall slowing of responding).

Discussion

This study provides provisional evidence that the emotional analogue of the Stroop paradigm may be sensitive to malingering. Although this study was limited by a small sample size, the group mean differences for interference scores were statistically significant nonetheless. The primary strengths of this study were the inclusion of a well-trained group of actors who were instructed by both doctoral-level psychologists and method acting coaches to feign a PTSD response and the inclusion of a PTSD trauma group that was not compensation seeking. Given the methodological difficulties associated with the study of malingering (see Freuh et al., 2000), the methodology employed by this study provided a novel examination of dissimulation.

The results revealed that the actor group was able to feign an overall delayed reaction time across all stimulus types that more closely resembled the PTSD group relative to the nonanxiety control group. However, the actors were unable to modulate their vocal response latency as a function of stimulus type in the manner that characterized the responses of the PTSD patients (the effect in the PTSD group was broadly consistent with what has been found in numerous other studies with PTSD samples; Buckley et al., 2000). The effect size estimates of these group differences were quite large and suggest that if we had a larger subject pool, we would have reached conventional levels of statistical significance to reject the null hypothesis on all comparisons.

Worthy of note is that our results might have been appreciably different if the actors had been coached on how to take the Stroop task as opposed to being coached only about symptomatic responses. However, we feel that our simulation most closely approximates what happens with malingering in real clinical practice and that coaching on the Stroop was contraindicated in that respect (i.e., people present to clinics with knowledge of symptomatic criteria, but without knowledge of specific response patterns on laboratory tasks or tests).

It has been argued that reaction-time-based information-processing tasks such as the Stroop may be harder to "fake" than face valid self-report instruments because the direction and magnitude of the disordered

response is not readily apparent to potential malingerers (Freuh et al., 2000). Our data were consistent with this notion; however, we had no data from self-report instruments with validity indices (e.g., the MMPI) to make such an inference from our study. Future studies may benefit from collecting both reaction time measures from information-processing tasks and standard measures of malingering like the validity indices of the MMPI for direct comparison purposes.

One might critique our study on the grounds that incentives to malinger might be much stronger in individuals who actually have a secondary-gain contingency that is operative (e.g., individuals who stand to gain financially should receive a diagnosis) relative to actors who are trained as confederates for a study in which the incentive value is lower. If present, such an effect would reduce the external validity of our findings to some extent in that the actors would not "feign" a PTSD response in the same way as actual malingerers. This is a valid concern for studies such as ours, however, malingerers, by definition, cannot be studied in a way that allows one to disentangle incentives, true responses, and malingered responses in clinical settings (Frueh et al., 2000). Doing so would require candid responding in the very environment in which those who are malingering have no incentive to do so! Thus, given such limitations in this research, we feel that our methods are a helpful first step in allowing us to examine the utility of information-processing measures as an index of malingering.

In light of the fact that information-processing indices may be sensitive to motivational levels, they should only be used in conjunction with other available data (as is the case with other measures of malingering). For example, Resnick (1997) points out that no medical test, psychophysiological test, or other index should be expected to "stand on its own" when trying to detect malingering. One reason is that motivation and other factors can adversely affect tests in ways that result in false positives or false negatives. So rather than relying on a single index that may be susceptible to such phenomenon, multiple sources of information should be utilized and the greatest point of convergence between measures should be used to arrive at a decision point. Thus, towards this end, reaction times to Stroop tasks may be useful *adjuncts* to clinicians who are attempting to discern psychopathology from malingering. In much the same way that it has been argued that psychophysiological measures can be useful *adjuncts* for diagnostic purposes (Blanchard & Buckley, 1999), we argue that information-processing indices might be used by clinicians in conjunction with other indices of malingering (e.g., psychophysiological response profiles, MMPI validity indices, neuropsychiatric tests of malingering). In this way, simple difference scores in reaction times (such as those analyzed in this paper) might be used as another point in a decision-making model. Future work with larger samples will be needed to determine what the appropriate cutoff scores (magnitude of differences) will be that will best assist with diagnostic decision-making in clinical practice.

To the extent that multiple sources of evidence are utilized, as outlined previously, and there is convergence on the same conclusion, the decision reached must be given more credence than one that was determined on the basis of a single index (Resnick, 1997). Given the ability of clinicians to administer Stroop tasks either on a computer, or more simply, with a standard card presentation (Williams et al., 1996), this is a resource that should be readily available to most clinicians to add to their clinical practice. However, given the important implications of rendering an opinion of either "malingering" or "PTSD" in cases where litigation is an issue, we advocate the use of computerized administration because the recording of response latencies is more accurate with that method relative to the card presentation (which is timed with a stop watch by a test administrator).

Finally we note that other information-processing paradigms might be even more useful than the Stroop for detecting malingering. We propose this in light of the aforementioned caveat that our results might have differed had the instructional set to the actors been different. The supposition that the results may have changed as a function of the instructional set implies that some aspect of the differences in reaction times are under volitional control. However, it has been argued that the vocal response delays during Stroop tasks seen in a variety of diagnostic groups are a function of primarily, automatic aspects of information processing (see McNally, 1995). Thus, to the extent that the modulation of vocal responses to various classes of stimuli is a function of processes that are beyond the volitional control of the patient, these effects should be extremely difficult to fake. Although our data are consistent with an "automaticity" hypothesis, there are numerous other methods for studying automatic processing of stimuli that have been applied to clinical samples (McNally, 1995). Moreover, many of the paradigms that evaluate automatic processing have been directly applied to PTSD populations (e.g., McNally, Amir, & Lipke, 1996). Thus, future studies may employ more stringent tests of whether or not reaction-time-based tasks can detect malingering by utilizing paradigms that tap more automatic aspects of information processing.

Although the detection of malingering has long been a problem for clinicians working with traumatized populations, the number of assessment tools to assist them is

still limited. Despite the fact that the data in this study were drawn from a small sample of subjects, the design was fairly strong and the findings suggest that the study of information-processing tasks as assessment devices for the detection of malingering warrants further investigation.

Acknowledgment

This research was supported by a grant from NIMH, MH-48476.

References

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 5,* 561–571.

Berry, D. T. R., Baer, R. A., & Harris, M. J. (1991). Detection of malingering on the MMPI: A meta-analytic review. *Clinical Psychology Review, 11,* 585–598.

Blair, J. R., & Spreen, O. (1989). Predicting premorbid IQ: A revision of the National Adult Reading Test. *Clinical-Neuropsychologist, 3,* 129–136.

Blake, D., Weathers, F., Nagy, L., Kaloupek, D., Klauminzer, G., Charney, D., et al. (1997). *Clinician-Administered PTSD Scale (CAPS).* Boston: National Center for Post-Traumatic Stress Disorder, Behavioral Science Division – Boston, VA.

Blanchard, E. B., & Buckley, T. C. (1999). Psychophysiological assessment of posttraumatic stress disorder. In P. A. Saigh & D. J. Bremner (Eds.), *Posttraumatic stress disorder: A comprehensive text* (pp. 248–266). Needham Heights, MA: Allyn & Bacon.

Blanchard, E. B., & Hickling, E. J. (1997). *After the crash.* Washington, DC: American Psychological Association.

Bryant, R. A., & Harvey, A. G. (1995). Processing threatening information in posttraumatic stress disorder. *Journal of Abnormal Psychology, 104,* 537–541.

Buckley, T. C., Blanchard, E. B., & Neill, W. T. (2000). Information processing and PTSD: A review of the empirical literature. *Clinical Psychology Review, 20,* 1041–1065.

Carroll, J. B., Davies, P., & Richman, B. (1971). *The American heritage word frequency book.* New York: American Heritage Publishing Company.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hilldale, NJ: Erlbaum.

Fairbank, J. A., McCaffrey, R. J., & Keane, T. M. (1985). Psychometric detection of fabricated symptoms of posttraumatic stress disorder. *American Journal of Psychiatry, 142,* 501–503.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1996). *Structured Clinical Interview for DSM-IV, Patient Edition (SCID-I/P)—Version 2.0.* New York: Biometrics Research Department – New York State Psychiatric Institute.

Foa, E. B., Feske, U., Murdock, T. B., Kozak, M. J., & McCarthy, P. R. (1991). Processing of threat-related information in rape victims. *Journal of Abnormal Psychology, 100,* 156–162.

Frueh, B. C., Hamner, M. B., Cahill, S. P., Gold, P. B., & Hamlin, K. (2000). Apparent symptom overreporting in combat veterans evaluated for PTSD. *Clinical Psychology Review, 20,* 853–885.

Frueh, B. C., & Kinder, B. N. (1994). The susceptibility of the Rorschach Inkblot Test to malingering of combat related PTSD. *Journal of Personality Assessment, 62,* 280–298.

Gerardi, R. J., Blanchard, E. B., & Kolb, L. C. (1989). Ability of Vietnam veterans to dissimulate a psychophysiological assessment for posttraumatic stress disorder. *Behavior Therapy, 20,* 229–243.

Harvey, A. G., Bryant, R. A., & Rapee, R. M. (1996). Pre-recognition processing of threat in posttraumatic stress disorder. *Cognitive Therapy and Research, 20,* 613–623.

Heiman, G. W. (1996). *Basic statistics for the behavioral sciences* (2nd ed.). Boston: Houghton Mifflin Company.

Hickling, E. J., Blanchard, E. B., Mundy, E., & Galovski, T. E. (2002). Detection of malingered MVA related posttraumatic stress disorder: An investigation of the ability to detect professional actors by experienced clinicians, psychological tests, and psychophysiological assessment. *Journal of Forensic Psychology Practice, 2,* 33–53.

Horowtiz, M. J., Wilmer, N., & Alvarez, N. (1979). Impact of event scale: A measure of subjective stress. *Psychosomatic Medicine, 41,* 209–218.

Lyons, J. A., Caddell, J. M, Pitman, R. K., Rawls, R., & Perrin, S. (1994). The potential for faking on the Mississippi Scale for combat-related PTSD. *Journal of Traumatic Stress, 7,* 441–445.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An intergrative review. *Psychological Bulletin, 109,* 163–203.

Mathews, A., & MacLeod, C. (1994). Cognitive approaches to emotion and emotional disorders. *Annual Review of Psychology, 45,* 25–50.

McNally, R. J. (1995). Automaticity and the anxiety disorders. *Behaviour Research and Therapy, 33,* 747–754.

McNally, R. J., Amir, N., & Lipke, H. J. (1996). Subliminal processing of threat cues in posttraumatic stress disorder? *Journal of Anxiety Disorders, 10,* 115–128.

Melton, G. B., Petrila, J., Poythress, N. G., & Slobogin, C. (1997). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers.* New York: Guilford Press.

Mogg, K., Kentish, J., & Bradley, B. P. (1993). Effects of anxiety and awareness on colour-identification latencies for emotional words. *Behaviour Research and Therapy, 31,* 559–567.

Morgan, C. A. (1997). Startle response in individuals with PTSD. *National Center for Posttraumatic Stress Disorder: Clinical Quarterly, 7,* 65–69.

Orr, S. P., & Pitman, R. K. (1993). Psychophysiologic assessment of attempts to simulate posttraumatic stress disorder. *Biological Psychiatry, 33,* 127–129.

Pitman, R. K., & Orr, S. P. (1993). Psychophysiologic testing for posttraumatic stress disorder: Forensic psychiatric application. *Bulletin of the American Academy of Psychiatry and Law, 21,* 37–52.

Resnick, P. J. (1997). Malingering of posttraumatic stress disorders. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 130–152). New York: Guilford Press.

Rothbaum, B. O., Foa, E. B., Riggs, D. S., Murdock, T., & Walsh, W. (1992). A prospective examination of post-traumatic stress disorder in rape victims. *Journal of Traumatic Stress, 5,* 455–475.

Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers. *Psychological Methods, 1,* 115–129.

Speilberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *STAI manual for the State-Trait Anxiety Inventory.* Palo Alto, CA: Consulting Psychologists Press.

Weathers, F. W., Litz, B. T., Herman, D. S., Huska, J. A., & Keane, T. M. (1993, October). *The PTSD checklist: Reliability, validity & diagnostic utility.* Paper presented at the Annual Meeting of the International Society for Traumatic Stress Studies, San Antonio, TX.

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin, 120,* 3–24.